JOURNAL OF DIFFERENTIAL EQUATIONS 4, 444–459 (1968)

On Differentials in Locally Convex Spaces

PETER L. FALB[1] AND MARC Q. JACOBS[2]

Center for Dynamical Systems, Brown University, Providence, Rhode Island 02912

Received April 20, 1967

1. INTRODUCTION

While developing some material related to direct methods in optimal control ([*1*]), we were led to the question of extending the notion of differential to the context of locally convex linear spaces. We answer this question here by defining suitable analogs of the strong (Frechet) and weak (Gateaux) differentials and by proving the standard results relating strong and weak differentiability [*2*]. Crucial to our treatment is the definition of $o(\cdot)$ for mappings between locally convex linear spaces.

The question of generalizing Frechet differentiability was considered by Michal [*3*] and Hyers [*4*]. However, our technique of proof differs from theirs and we prove both a theorem relating Gateaux to Frechet differentiability (Theorem 2) and an implicit function theorem (Theorem 3), neither of which appear in their work. Moreover, the implicit function theorem can be used to obtain a generalization of the Lagrange multiplier rule in minimization problems (see [*1*] and appendix 2).

2. DEFINITIONS

Let E be a locally convex topological vector space.[3] We recall that the topology on E is determined by the family of all continuous seminorms on E [*5*]. Any family Q_E of continuous seminorms on E which determines the

[1] This research was supported in part by the National Science Foundation, under Grant GK-967 and in part by the Air Force Office of Scientific Research, Office of Aerospace Research, U. S. Air Force under AFOSR Grant 693-67.

[2] This research was supported in part by the National Science Foundation, under Grant GP-5970, in part by the Air Force Office of Scientific Research, Office of Aerospace Research, U. S. Air Force under AFOSR Grant 693-67 and in part by the United States Army Research Office, Durham, under Contract DA-31-124-ARO-D-270.

[3] All the spaces that we shall consider are assumed to be Hausdorff spaces.

topology on E shall be called a *generating family*. We use the notation Q_E to indicate a generating family of seminorms throughout the sequel.

Now let F be another locally convex linear space and let ψ be a mapping of E into F. We then have

DEFINITION 1. The mapping ψ is said to be $o(h)$ if there is a generating family Q_F such that "$q \in Q_F$ and $\epsilon > 0$" implies that there is a continuous seminorm p on E (which may depend on q and ϵ) and a $\delta > 0$ such that $p(h) < \delta$ implies that $q(\psi(h)) \leqslant \epsilon p(h)$.

The following proposition is a statement of the independence of Definition 1 of the generating family Q_F:

PROPOSITION 1. *Let ψ be $o(h)$ with respect to Q_F and let Q'_F be another generating family. Then ψ is $o(h)$ with respect to Q'_F.*

Proof. We may assume, without loss of generality, that Q_F is saturated [5]. Since the identity map is continuous, $q' \in Q'_F$ implies that there is a q in Q_F and an $M > 0$ such that $q'(x) \leqslant Mq(x)$ for all x in F ([5], p. 97). However, if $\epsilon > 0$, then there is a p and a $\delta > 0$ such that $p(h) < \delta$ implies $q(\psi(h)) \leqslant \epsilon p(h)/M$ as ψ is $o(h)$ with respect to Q_F. It follows that $q'(\psi(h)) \leqslant \epsilon p(h)$ and so the proposition is established.

We next show that Definition 1 agrees with the usual definition for normed linear spaces.

PROPOSITION 2. *If E and F are normed, then ψ is $o(h)$ in the sense of Definition 1 if and only if ψ is $o(\| h \|)$.*

Proof. If ψ is $o(\| h \|)$, then we simply take $Q_F = \{\| \cdot \|_F\}$. On the other hand, suppose that ψ is $o(h)$. We first note that if p is a continuous seminorm on E, then there is an $M_p > 0$ such that $p(\cdot) \leqslant M_p \| \cdot \|_E$. Thus, in view of Proposition 1, $\epsilon > 0$ implies that there is a continuous seminorm p on E and a $\delta > 0$ such that $\| \psi(h)\|_F \leqslant \epsilon p(h)/M_p \leqslant \epsilon \| h \|_E$ if $p(h) < \delta$. It follows that if $\| h \|_E < \delta/M_p$, then $\| \psi(h)\|_F \leqslant \epsilon \| h \|_E$ which establishes the proposition.

With the notion of $o(h)$ in hand, we have the following:

DEFINITION 2. Let f be a mapping of E into F. *If* x is an element of E and if there is a continuous linear transformation $df(x, \cdot)$ of E into F such that

$$f(x + h) - f(x) = df(x, h) + \psi(x, h)$$

where $\psi(x, \cdot)$ is $o(h)$, then f is strongly (Frechet) differentiable at x and $df(x, h)$ is the strong differential of f at x in the direction h.

DEFINITION 3. Let f be a mapping of E into F. If x is an element of E and if the limit

$$\lim_{t\to 0} \frac{f(x+th)-f(x)}{t} = \frac{d}{dt} f(x+th)\Big|_{t=0}$$

exists, then this limit is called the weak (Gateaux) differential of f at x in the direction h and is denoted by $d_w f(x, h)$. f is weakly differentiable at x if $d_w f(x, h)$ exists for all h in E.

We prove two theorems relating these notions of differentiability in the next section.

3. Two Theorems

THEOREM 1. *If f is strongly differentiable at x, then f is weakly differentiable at x and $d_w f(x, h) = df(x, h)$ for all h in E.*

Proof. We have

$$f(x+th) - f(x) = t\, df(x, h) + \psi(x, th)$$

where $\psi(x, \cdot)$ is $o(h)$. Thus it will be enough to show that

$$\lim_{t\to 0} \frac{\psi(x, th)}{t} = 0.$$

To do this, we let Q_F be a generating family for F. Then $\epsilon > 0$ and $q \in Q_F$ imply that there is a continuous seminorm p on E and a $\delta > 0$ such that $q(\psi(x, th)) \leqslant \epsilon p(th) = \epsilon \mid t \mid p(h)$ whenever $p(th) < \delta$. It follows that $\mid q(\psi(x, th))/t \mid \leqslant \epsilon p(h)$ and hence that $\lim_{t\to 0} q(\psi(x, th))/t = 0$. As q is an arbitrary element of Q_F, the theorem is established.

The next theorem, which is the analog of Theorem 2, p. 185, of [*2*], requires the notion of Riemann integrability for functions taking values in a complete locally convex linear space. We develop the relevant results in Appendix 1.

THEOREM 2. *Let F be a complete locally convex linear space. Let U be a closed, convex, balanced and absorbing neighborhood of 0 in E and let $D = x_0 + U$ where x_0 is a given element of E. If f is weakly differentiable on D, if $d_w f(x, \cdot)$ is a continuous map of E into F for all x in D, and if $d_w f(\cdot, h)$ is a uniformly-continuous mapping (with respect to h in U) at each x in D, then f is strongly differentiable at every interior point y of D and $df(y, h) = d_w f(y, h)$ for all h in E.*

Proof. It will be sufficient to consider the point x_0 only as the arguments are essentially unchanged for any other interior point of D. Now if $h \in U$, then $x_{th} = x_0 + th$, $0 \leqslant t \leqslant 1$, is in U and

$$d_w f(x_{th}, h) = \lim_{s \to 0} [f(x_{th} + sh) - f(x_{th})]/s$$

exists. Since $x_{th} + sh = x_0 + (t + s)h$, it follows that

$$d_w f(x_{th}, h) = \frac{d}{dt} f(x_0 + th). \tag{1}$$

We now claim that $d_w f(x_0, \cdot)$ is additive, i.e., that

$$d_w f(x_0, h_1 + h_2) = d_w f(x_0, h_1) + d_w f(x_0, h_2) \tag{2}$$

for all h_1, h_2 in E. Assuming that (2) is valid, we immediately deduce that $d_w f(x_0, \cdot)$ is a continuous *linear* transformation since $d_w f(x_0, \cdot)$ is continuous and additive.

Let us establish (2). Since U is convex, balanced and absorbing, there is a t_0 with $0 < t_0 \leqslant 1$ such that $0 \leqslant t \leqslant t_0$ implies that th_1, th_2 and $t(h_1 + h_2)$ are all in U. It follows from (1) and the properties of the Riemann integral (see Appendix 1) that

$$f(x_0 + th_1) - f(x_0) = \int_0^t d_w f(x_0 + \xi h_1, h_1)\, d\xi = \alpha_1 + t d_w f(x_0, h_1) \tag{3a}$$

and similarly that

$$f(x_0 + t(h_1 + h_2)) - f(x_0) = \alpha_2 + t\, d_w f(x_0, h_1 + h_2) \tag{3b}$$

$$f(x_0 + t(h_1 + h_2)) - f(x_0 + th_1) = \alpha_3 + t\, d_w f(x_0, h_2) \tag{3c}$$

for $0 \leqslant t \leqslant t_0$, where

$$\begin{aligned} \alpha_1 &= \int_0^t \{d_w f(x_0 + \xi h_1, h_1) - d_w f(x_0, h_1)\}\, d\xi \\ \alpha_2 &= \int_0^t \{d_w f(x_0 + \xi(h_1 + h_2), h_1 + h_2) - d_w f(x_0, h_1 + h_2)\}\, d\xi \\ \alpha_3 &= \int_0^t \{d_w f(x_0 + th_1 + \xi h_2, h_2) - d_w f(x_0, h_2)\}\, d\xi. \end{aligned} \tag{4}$$

Now let N be any neighborhood of 0 in F and let K be a closed, convex, symmetric neighborhood of 0 in F such that $K + K + K \subset N$. By the continuity of $d_w f(\cdot, h)$, there is a t_1 with $0 \leqslant t_1 \leqslant t_0$ such that the integrands in (4) are all in K for $0 \leqslant \xi \leqslant t_1$. Since K is closed, convex and symmetric,

$\pm\alpha_i \in tK, i = 1, 2, 3$ whenever $0 \leqslant t \leqslant t_1$. Thus, we deduce from (3) that

$$t\{d_w f(x_0, h_1) + d_w f(x_0, h_2) - d_w f(x_0, h_1 + h_2)\} + (\alpha_1 + \alpha_3 - \alpha_2) = 0$$

and consequently, that $d_w f(x_0, h_1) + d_w f(x_0, h_2) - d_w f(x_0, h_1 + h_2)$ is in N. As N was any neighborhood of 0 in F, (2) is established.

All that remains is to show that the mapping $\psi(x_0, \cdot)$ given by

$$f(x_0 + h) - f(x_0) = d_w f(x_0, h) + \psi(x_0, h)$$

is $o(h)$. If $h \in U$, then

$$f(x_0 + h) - f(x_0) = \int_0^1 f'(x_0 + th)h\,dt = \alpha + f'(x_0)h$$

where $f'(x)\,h = d_w f(x, h)$ and

$$\alpha = \int_0^1 \{f'(x_0 + th)h - f'(x_0)h\}\,dt. \tag{5}$$

Letting Q_F be a generating family and noting that $f'(\cdot)\,h$ is continuous at x_0 (uniformly on U), we have "$q \in Q_F$ and $\epsilon > 0$" implies that there is a continuous seminorm p on E such that for all u in U,

$$q(\{f'(x_0 + th) - f'(x_0)\}\,u) \leqslant \epsilon$$

whenever $p(h) < 1$. Letting g_U be the gauge of U and noting that $p^* = \max(p, g_U)$ is a continuous seminorm on E, we have

$$q(\{f'(x_0 + th) - f'(x_0)\}\,u) \leqslant \epsilon$$

whenever $p^*(h) < 1$ and $p^*(u) < 1$. Now let e be any element of E and let s be *any* real number such that $p^*(e) < s$. Then $p^*(e/s) < 1$ and $q(\{f'(x_0 + th) - f'(x_0)\}\,e) \leqslant \epsilon s$ if $p^*(h) < 1$. It follows that $p^*(h) < 1$ implies that

$$q(\{f'(x_0 + th) - f'(x_0)\}\,e) \leqslant \epsilon p^*(e) \tag{6}$$

for all e in E. Combining (5) and (6), we deduce that

$$\begin{aligned} q(\psi(x_0, h)) &= q\left(\int_0^1 \{f'(x_0 + th) - f'(x_0)\}h\,dt\right) \\ &= \int_0^1 q(\{f'(x_0 + th) - f'(x_0)\}h)\,dt \\ &\leqslant \epsilon p^*(h) \end{aligned}$$

whenever $p^*(h) < 1$, since q is continuous. Thus the proof of Theorem 2 is complete.

4. Three Simple Propositions

Let E, F and G be locally convex topological vector spaces. We now give three simple propositions which will be used in our treatment of the implicit function theorem in Section 5.

Proposition 3. *Let T be a continuous linear transformation of E into F and let Q_F be a generating family of seminorms on F. If $q \in Q_F$, then there is a continuous seminorm p on E such that $q(T(x)) \leqslant p(x)$ for all x in E.*

Proposition 4. *Let ψ_1 and ψ_2 be mappings of E into F and of F into G, respectively. If ψ_1 is $o(s)$ and ψ_2 is $o(t)$, then $\psi_2 o \psi_1$ is $o(s)$.*

Proof. Let Q_G be a generating family for G and let Q_F be the family of *all* continuous seminorms on F. Now, $q \in Q_G$ implies there is an r in Q_F and a $\delta_r > 0$ such that $q(\psi_2(y)) \leqslant r(y)$ if $r(y) < \delta_r$. Similarly, if $0 < \epsilon < 1$, then there is a continuous seminorm p on E and a δ_p with $0 < \delta_p < \delta_r$ such that $r(\psi_1(x)) \leqslant \epsilon p(x)$ if $p(x) < \delta_p$. Thus, $p(x) < \delta_p$ implies that $r(\psi_1(x)) < \delta_r$ and hence that $q(\psi_2(\psi_1(x))) \leqslant r(\psi_1(x)) \leqslant \epsilon p(x)$.

Proposition 5. *Let A and B be open sets in E and F, respectively; let f be a Frechet differentiable map of A into B; and, let g be a Frechet differentiable map of B into G. If x_0 is an element of A, then $h = gof$ is Frechet differentiable at x_0 and*

$$dh(x_0, s) = dg(f(x_0), df(x_0, s)). \tag{7}$$

Proof. Simply apply Definition 2 and Propositions 3 and 4.

5. An Implicit Function Theorem

Before proving the theorem, we make the following definitions:

Definition 4. A subset U of a topological vector space E is a full domain if (i) there is a family $\{U_\alpha\}$ of convex open sets in E such that $U = \cap U_\alpha$, and (ii) U contains a simplex of dimension greater than zero.

Definition 5. Let E and F be locally convex topological vector spaces with generating families Q_E and Q_F, respectively. Let U be a full domain in

E and let f be a mapping of U into F. Then f is differentiable at an element x_0 of U if there is a linear mapping $df(x_0, \cdot)$ of E into F such that (i) $df(x_0, \cdot)$ restricted to U is continuous, and (ii) "$q \in Q_F$ and $\epsilon > 0$" imply that there is a p in Q_E and a $\delta > 0$ with the property that

$$q(f(x_0 + h) - f(x_0) - df(x_0, h)) \leqslant \epsilon p(h) \tag{8}$$

if $x_0 + h \in U$ and $p(h) < \delta$.

We let E, F and G be locally convex topological vector spaces and we now enumerate a number of conditions which will be used in Theorem 3:

(c1) F is sequentially complete.

(c2) A is an open set in $E \times F$.

(c3) f is a mapping of A into G and (x_0, y_0) is an element of A for which $f(x_0, y_0) = 0$.

(c4) f is Frechet differentiable at (x_0, y_0).

(c5) U_0 is a full domain in E and x_0 is an element of U_0.

(c6) $u(\cdot)$ is a continuous mapping of U_0 into F with the following properties: (i) $u(x_0) = y_0$, (ii) $(x, u(x)) \in A$ for all x in U_0, and (iii) $f(x, u(x)) = 0$ for all x in U_0.

(c7) The partial derivative $\partial f(x_0, y_0)/\partial y$ is a linear homeomorphism of F onto G.

(c8) There is a generating family Q_F on F such that $q \in Q_F$ and $\epsilon > 0$ imply that there is a $\delta > 0$ for which

$$q\left(\left[\frac{\partial f}{\partial y}(x_0, y_0)\right]^{-1} \left\{\frac{\partial f}{\partial y}(x_0, y_0)k - [f(x_0 + h, y_0 + k) - f(x_0 + h, y_0)]\right\}\right) \leqslant \epsilon q(k)$$

if $q(k) < \delta$ and $(x_0 + h, y_0 + k) \in A$.

(c8*) There are, (i) a generating family Q_F on F, (ii) an $\epsilon \leqslant \frac{1}{2}$, (iii) a full domain U_ϵ in E, and (iv) a bounded, closed full domain V_ϵ in F, such that the following conditions are satisfied:

(a) $(x_0, y_0) \in U_\epsilon \times V_\epsilon \subset A$;

(b) $V_\epsilon = \cap V_i$ where $V_i = y_0 + \{y: q_i(y) < \epsilon_i\}$ and the q_i are seminorms in Q_F;

(c) $f(U_\epsilon, y_0) \subset \partial f/\partial y(x_0, y_0)\, \tilde{V}_\epsilon$ where $\tilde{V}_\epsilon = \cap \tilde{V}_i$ with $\tilde{V}_i = y_0 + \epsilon\{y: q_i(y) < \epsilon_i\}$; and,

(d) if $x \in U_\epsilon$, if y_1, y_2 are in V_ϵ and if $q \in Q_F$, then

$$q\left(\left[\frac{\partial f}{\partial y}(x_0, y_0)\right]^{-1} \left\{\frac{\partial f}{\partial y}(x_0, y_0)(y_1 - y_2) - [f(x, y_1) - f(x, y_2)]\right\}\right) \leqslant \epsilon q(y_1 - y_2).$$

We now have

THEOREM 3. (A) *Suppose that* (c2)–(c8) *are satisfied. Then* $u(\cdot)$ *is Frechet differentiable at the element* x_0 *of* U_0 *and*

$$du(x_0, h) = -\left[\frac{\partial f}{\partial y}(x_0, y_0)\right]^{-1} \frac{\partial f}{\partial x}(x_0, y_0)h$$

for h *in* U_0. (B) *Suppose that* (c1)–(c4), (c7) *and* (c8*) *are satisfied. Then there is a full domain* U_0 *in* E *with* x_0 *in* U_0 *and a continuous mapping* $u(\cdot)$ *of* U_0 *into* F *for which* (c6) *is satisfied. Moreover, if* (c8) *is satisfied, then* $u(\cdot)$ *is Frechet differentiable at* x_0 *and*

$$du(x_0, h) = -\left[\frac{\partial f}{\partial y}(x_0, y_0)\right]^{-1} \frac{\partial f}{\partial x}(x_0, y_0)h$$

for h *in* U_0.

Proof. We suppose, without loss of generality, that $(x_0, y_0) = (0, 0)$ and we let $S_0 = \partial f/\partial x(0, 0)$, $T_0 = \partial f/\partial y(0, 0)$.

Let us first prove (A). We note that $u(h) \to 0$ as $h \to 0$ and that the mapping $\psi(h)$ given by

$$\psi(h) = f(h, 0) - f(0, 0) - S_0(h) \tag{9}$$

is $o(h)$ as f is differentiable at (0, 0). Moreover,

$$-[S_0(h) + T_0(u(h))] = [f(h, u(h)) - f(h) - T_0(u(h))] + \psi(h) \tag{10}$$

as $f(h, u(h)) = 0$ for h in U_0. Now $\epsilon > 0$ and $q \in Q_F$ imply, by virtue of (c8), that there is a $\delta_1 > 0$ such that

$$q(T_0^{-1}\{f(h, u(h)) - f(h, 0) - T_0(u(h))\}) \leqslant \epsilon q(u(h)) \tag{11}$$

if $h \in U_0$ and $q(u(h)) < \delta_1$. Since f is Frechet differentiable at (0, 0) and since $\{qoT_0^{-1}\colon q \in Q_F\}$ is a generating family for G, there is a continuous seminorm p_1 on E and there is a δ_2 with $0 < \delta_2 < 1$ such that

$$q(T_0^{-1}\{\psi(h)\}) \leqslant \epsilon p_1(h) \tag{12}$$

if $p_1(h) < \delta_2$. As $u(\cdot)$ is continuous on U_0, there is a continuous seminorm p_2 on E such that $h \in U_0$ and $p_2(h) < 1$ together imply that $q(u(h)) < \delta_1$. Let $p_3 = \max(p_1, p_2)$ and $\delta_3 = \min(\delta_1, \delta_2)$. Then (10), (11) and (12) yield the inequality

$$q(u(h) + T_0^{-1}S_0(h)) \leqslant \epsilon(p_3(h) + q(u(h))) \tag{13}$$

when $h \in U_0$ and $p_3(h) < \delta_3$. Since $T_0^{-1}S_0$ is continuous, there is a continuous seminorm p_4 on E such that

$$q(T_0^{-1}S_0(h)) \leqslant p_4(h) \tag{14}$$

for all h in E. Note that p_4 does not depend on ϵ. Combining (13) and (14), we find that

$$q(u(h)) \leqslant p_4(h) + \epsilon(p_3(h) + q(u(h))) \tag{15}$$

if $h \in U_0$ and $p_3(h) < \delta_3$. We note that p_3 and δ_3 depend on ϵ. In *particular*, for $\epsilon = \hat{\epsilon} = \frac{1}{2}$, we deduce that there is a $\hat{\delta}_3$ (and a $\hat{p}_3$) such that

$$q(u(h)) \leqslant 2p_4(h) + \hat{p}_3(h) \tag{16}$$

if $h \in U_0$ and $\hat{p}_3(h) < \hat{\delta}_3$. Now let $p = \max(p_4, p_3, \hat{p}_3)$ and let $\delta = \min(\hat{\delta}_3, \delta_3)$. Then, combining (13) and (16), we have

$$q(u(h) + T_0^{-1}S_0(h)) \leqslant 4\epsilon p(h) \tag{17}$$

whenever $h \in U_0$ and $p(h) < \delta$. Thus $u(\cdot)$ is differentiable at the element 0 of U_0 and

$$du(0, h) = -T_0^{-1}S_0(h)$$

for h in U_0. Thus (A) is established.

Now let us prove (B). Letting $g(x, y)$ be the mapping of A into F given by

$$g(x, y) = y - T_0^{-1}f(x, y), \tag{18}$$

we observe that $f(x, y) = 0$ if and only if $g(x, y) = y$. If $q \in Q_F$, $x \in U_\epsilon$ and $y_1, y_2 \in V_\epsilon$, then

$$q(g(x, y_1) - g(x, y_2)) \leqslant \epsilon q(y_1 - y_2) \tag{19}$$

by (d) of (c8*). For $x \in U_0$, we set $u_0(x) = 0$ and $u_1(x) = g(x, 0)$. In view of (c) of (c8*), $u_1(x)$ is in $\tilde{V}_\epsilon \subset V_\epsilon$. Now let us suppose that $u_n(x) = g(x, u_{n-1}(x))$ is defined and is an element of V_ϵ for $1 \leqslant n \leqslant m$. Then, $u_{m+1}(x) = g(x, u_m(x))$ is defined and

$$q(u_{m+1}(x) - u_m(x)) \leqslant \epsilon q(u_m(x) - u_{m-1}(x)) \leqslant \epsilon^m q(u_1(x)) \tag{20}$$

for all q in Q_F in view of (19). In particular, (20) implies that

$$q_i(u_{m+1}(x)) \leqslant (1 + \epsilon + \cdots + \epsilon^m)\, q_i(u_1(x)) < \frac{1}{1-\epsilon}\, q_i(u_1(x)) \tag{21}$$

for *all* the seminorms q_i defining V_ϵ. Since $u_1(x)$ is in $\tilde{V}_\epsilon$, $q_i(u_1(x)) < \epsilon\epsilon_i$ for all i and so,

$$q_i(u_{m+1}(x)) < \frac{\epsilon}{1-\epsilon} \cdot \epsilon_i \leqslant \epsilon_i \tag{22}$$

as $\epsilon \leqslant \frac{1}{2}$. Thus, $u_{m+1}(x)$ is in V_ϵ. It follows, by induction, that $u_n(x) = g(x, u_{n-1}(x))$ is in V_ϵ for every n. In view of (20), $\{u_n(x)\}$ is a Cauchy sequence in F for every q in Q_F. As F is sequentially complete and V_ϵ is closed, $\{u_n(x)\}$ converges to a unique element $u(x)$ of V_ϵ. Thus, we have defined a mapping $u(\cdot)$ of U_ϵ into V_ϵ such that $u(0) = 0$, $(x, u(x)) \in A$ and $f(x, u(x)) = 0$ for all x in U_ϵ. Since V_ϵ is bounded, (20) implies that the sequence of continuous functions $u_n(x)$ converges to $u(x)$ uniformly on U_ϵ and, therefore, $u(x)$ is continuous on U_ϵ. Thus, part (B) is established.

Examples to which this theorem applies are easily constructed using product spaces. For example, let $E = F = R^R$ be the R-fold topological product of R with itself (where R has the usual topology). Elements of R^R are written in the form $\{x_t\}$ where $x_t \in R$, $t \in R$. Consider the mapping f of $R^R \times R^R$ into R^R given by $f(\{x_t\}, \{y_t\}) = \{e^{y_t} - (x_t + 1)\}$. Then $f(\{0_t\}, \{0_t\}) = \{0_t\}$ and the hypotheses of Theorem 3 can easily be verified.

Appendix 1

The Riemann Integral in a Complete Locally Convex Linear Space

Let $[a, b]$ denote a compact interval in R. By a *partition* of $[a, b]$ we shall mean a system $\pi = (I_1, \ldots, I_k; \xi_1, \ldots, \xi_k)$ consisting of disjoint intervals $I_j, j = 1, 2, \ldots, k$, with $\bigcup I_j = [a, b]$ and points $\xi_j, j = 1, 2, \ldots, k$ such that $\xi_j \in I_j, j = 1, 2, \ldots, k$. Let $\mathscr{P}$ denote the totality of all partitions of the interval $[a, b]$. We say that $\pi_2 = (I_1, \ldots, I_k; \xi_1, \ldots, \xi_k)$ is a *refinement* of $\pi_1 = (J_1, \ldots, J_p; \eta_1, \ldots, \eta_p)$ if each $J_i, i = 1, 2, \ldots, p$ is expressible as the union of some of the $I_m, m = 1, 2, \ldots, k$. A relation $\geqslant$ is defined on $\mathscr{P}$ by defining $\pi_2 \geqslant \pi_1$ if and only if π_2 is a refinement of π_1. Then $(\mathscr{P}, \geqslant)$ is a directed set.

Let F be a complete locally convex linear space (Hausdorff). Let x be a continuous map of $[a, b]$ into F. A mapping S_x of $\mathscr{P}$ into F is defined as follows:

$$S_x(\pi) = \sum_{j=1}^{k} x(\xi_j)\, \lambda(I_j),$$

where $\pi = (I_1, \ldots, I_k; \xi_1, \ldots, \xi_k)$ and $\lambda(I)$ denotes the length of an interval. Thus, $(S_x, \mathscr{P}, \geqslant)$ is a net in F. If the net $(S_x, \mathscr{P}, \geqslant)$ converges to a limit

h, then x is said to be *Riemann integrable* on $[a, b]$, and h is called the *Riemann integral of x on* $[a, b]$ and is denoted by $\int_a^b x(t)\,dt$.

The following are basic properties of the Riemann integral:

PROPERTY 1. *A continuous function x from $[a, b]$ into F is Riemann integrable.*

Proof. We show that $(S_x, \mathscr{P}, \geqslant)$ is a Cauchy net. Let U denote an arbitrary convex neighborhood of 0 in F. Then there is a V, which is a balanced, convex neighborhood of 0 such that $V + V \subset U$. Since x is continuous on the compact interval $[a, b]$, x is uniformly continuous on $[a, b]$. Consequently there is a $\delta > 0$ such that $t, t' \in [a, b]$ and $|t - t'| < \delta$ imply $x(t) - x(t') \in V$. Choose a partition $\pi_0 = (I_1{}^0,..., I_N{}^0; \xi_1{}^0,..., \xi_N{}^0)$ such that $\max_j \lambda(I_j) \leqslant \delta$. Let $\pi_1 = (I_1{}^1,..., I_{N1}^1 ; \xi_1{}^1,..., \xi_{N1}^1)$ and $\pi_2 = (I_1{}^2,..., I_{N2}^2 ; \xi_1{}^2,..., \xi_{N2}^2)$ be refinements of π_0. Since V is balanced and convex, and since $\pi_1, \pi_2 \geqslant \pi_0$, we have

$$[S_x(\pi_1) - S_x(\pi_0)], [S_x(\pi_0) - S_x(\pi_2)] \in (b - a)\, V.$$

It follows that $S_x(\pi_1) - S_x(\pi_2)$ belongs to $(b - a)(V + V) \subset (b - a)\, U$, and hence that $(S_x, \mathscr{P}, \geqslant)$ is a Cauchy net. As F is complete, the property holds.

The next two properties are evident, so we shall omit their proof.

PROPERTY 2. *Let x, y be continuous mappings of $[a, b]$ into F. Let f be a continuous linear functional on F and let q be a continuous seminorm on F. Then*

$$\text{(i)} \quad \int_a^b [\alpha x(t) + \beta y(t))]\, dt = \alpha \int_a^b x(t)\, dt + \beta \int_a^b y(t)\, dt, \alpha, \beta \in R;$$

$$\text{(ii)} \quad f\left(\int_a^b x(t)\, dt\right) = \int_a^b f(x(t))\, dt;$$

$$\text{(iii)} \quad q\left(\int_a^b x(t)\, dt\right) \leqslant \int_a^b q(x(t))\, dt.$$

PROPERTY 3. *Let x be a continuous map of $[a, b]$ into F and let K be a closed convex subset of E. If $x(t)$ is an element of K for each t in $[a, b]$, then $\int_a^b x(t)\, dt$ is an element of $(b - a)\, K$.*

PROPERTY 4. *If x_n, $n = 1, 2, 3,...$, is a sequence of continuous functions from $[a, b]$ into F such that x_n converges to 0 uniformly as $n \to \infty$, then* $\lim_{n\to\infty} \int_a^b x_n(t)\, dt = 0$.

Proof. Let U be a closed, convex neighborhood of 0 in F. Then there is an n_0 such that $x_n(t) \in U$, $a \leqslant t \leqslant b$, for all $n \geqslant n_0$. By Property 3, $\int_a^b x_n(t)\, dt \in (b-a)\, U$ whenever $n \geqslant n_0$. Therefore $\int_a^b x_n(t)\, dt$ approaches 0 as $n \to \infty$.

PROPERTY 5. *Let x be a continuous mapping of $[a, b]$ into F and let $G(t) = \int_a^t x(\xi)\, d\xi$ for t in $[a, b]$. Then $dG(t)/dt$ exists and is equal to $x(t)$ for t in $[a, b]$.*

Proof. Let U be a closed convex neighborhood of 0 in F. Then

$$\frac{\Delta G}{\Delta t} = \frac{G(t+\Delta t) - G(t)}{\Delta t} = \frac{1}{\Delta t}\int_t^{t+\Delta t} x(\xi)\, d\xi.$$

Since x is continuous at t, there is a $\delta > 0$ such that $x(\xi) \in x(t) + U$, whenever $|\xi - t| \leqslant \delta$ and $\xi \in [a, b]$. Therefore, if $|\Delta t| \leqslant \delta$ and $t + \Delta t \in [a, b]$, the mapping x restricted to $[t, t + \Delta t]$ has its range in $x(t) + U$ and consequently we have

$$\frac{1}{\Delta t}\int_t^{t+\Delta t} x(\xi)\, d\xi = \frac{\Delta G}{\Delta t} \in x(t) + U.$$

It follows that $\lim_{\Delta t \to 0} \Delta G/\Delta t = x(t)$.

PROPERTY 6. [6] *Let x be a differentiable mapping of $[a, b]$ into F. Let q be a continuous seminorm on F and let M be a real number such that $q(x'(t)) \leqslant M$ (where $x'(t) = dx(t)/dt$) for t in $[a, b]$. Then $q(x(b) - x(a)) \leqslant M(b-a)$.*

Proof. Let $\epsilon > 0$ be given and let A_ϵ be the subset of $[a, b]$ defined by the condition that ξ is in A_ϵ if and only if $a \leqslant \eta < \xi$ implies that

$$q(x(\eta) - x(a)) \leqslant M(\eta - a) + \epsilon(\eta - a).$$

Evidently a is in A_ϵ. Moreover, if ξ is in A_ϵ and $a < \eta < \xi$, then η is in A_ϵ. Let ω denote the supremum of A_ϵ. It follows from the definition of A_ϵ that A_ϵ is the compact interval $[a, \omega]$. Since x is continuous at ω,

$$q(x(\omega) - x(a)) \leqslant M(\omega - a) + \epsilon(\omega - a).$$

Since $\epsilon > 0$ is arbitrary, we need only show that $\omega = b$. Suppose that $\omega < b$. Then there is an interval $[\omega, \omega + \delta]$ contained in $[a, b]$ ($\delta > 0$) such that $\omega \leqslant t < \omega + \delta$ implies that

$$\begin{aligned} |\, q(x(t) - x(\omega)) - q(x'(\omega)(t-\omega))| &\leqslant q(x(t) - x(\omega) - x'(\omega)(t-\omega)) \\ &\leqslant \epsilon(t - \omega). \end{aligned}$$

Consequently, if $a \leqslant t < \omega + \delta$, then

$$q(x(t) - x(\omega)) \leqslant (t - \omega)\, q(x'(\omega)) + \epsilon(t - \omega) \leqslant M(t - \omega) + \epsilon(t - \omega),$$

which implies $\omega + \delta$ is in A_ϵ, contrary to the definition of ω.

PROPERTY 7. *If x is a differentiable mapping of $[a, b]$ into F and if $x'(t) = 0$ for each t, then $x(t) = x(a)$ for all t in $[a, b]$.*

Proof. Let $M = 0$. Then the conditions of Property 6 are satisfied for every continuous seminorm q on F. Since F is Hausdorff, the conclusion follows immediately.

PROPERTY 8. *Let x be a continuously differentiable mapping of $[a, b]$ into F; then*

$$\int_a^b x'(t)\, dt = x(b) - x(a).$$

Proof. An immediate consequence of Properties 5 and 7.

APPENDIX 2

A Lagrange Multiplier Rule

Let E and H be locally convex linear spaces and let g be a mapping of H into E. We then have

DEFINITION. If $x \in H$, then x is called regular for g if the following conditions are satisfied:

(i) $dg(x, \cdot) = g'(x)$ exists;

(ii) the kernel of $g'(x)$, $K_x(g) = \{h : g'(x)\, h = 0_E\}$, has a topological supplement $\hat{K}_x(g)$ in H; and,

(iii) $g'(x)$ is a surjective strict morphism [5]; i.e., $g'(x)$ when restricted to $\hat{K}_x(g)$ is a linear homeomorphism between $\hat{K}_x(g)$ and E.

We let $S(g)$ denote the set of zeroes of g; i.e., $S(g) = \{h : g(h) = 0_E\}$. We say that $S(g)$ is *smooth at h* if $h \in S(g)$ and h is regular for g. If $S(g)$ is smooth at all of its points, then we call $S(g)$ a "*smooth*" *variety.*

Now let us suppose that x is an element of H which is regular for g. We then have

DEFINITION. If H and E are normed, then x is called admissible for g if

$dg(\cdot, \cdot) = g'(\cdot)$ is continuous on a neighborhood $\mathbf{N} + x$ of x (when viewed as a map of H into $\mathscr{L}(H, E)$).

If H and E are normed and x is an element of $S(g)$, then we shall also require that $\mathbf{N} \cap S(g)$ consist of regular points for g as part of the notion that x is admissible for g.

Definition. If either H or E is not normed, then x is called admissible for g if there is a neighborhood $\mathbf{N}$ of 0_H and a generating family $\hat{Q}_x$ on $\hat{K}_x(g)$ such that the following conditions are satisfied:

(a) If h is an element of $K_x(g)$, then there are an $\epsilon_0 \leqslant \frac{1}{2}$, a bounded closed full domain $V_{\epsilon_0} \subset \hat{K}_x(g)$, and an interval $U_{\epsilon_0} \subset R$ with $0 \in \operatorname{int}(U_{\epsilon_0})$, ($\epsilon_0$, V_{ϵ_0} and U_{ϵ_0} may depend on h) such that

(i) $0_H \in U_{\epsilon_0} h \oplus V_{\epsilon_0} \subset \mathbf{N}$,

(ii) $V_{\epsilon_0} = \cap V_i$ with $V_i = \{y \in \hat{K}_x(g) \colon \hat{q}_i(y) < \epsilon_i\}$ where the $\hat{q}_i$ are in $\hat{Q}_x$,

(iii) $g(x + U_{\epsilon_0} h) \subset g(x) + g'(x)\, \tilde{V}_{\epsilon_0}$ where $\tilde{V}_{\epsilon_0} = \cap \tilde{V}_i$ with $\tilde{V}_i = \epsilon_0 V_i$,

(iv) if $z \in U_{\epsilon_0} h$, $y_1, y_2 \in V_{\epsilon_0}$ and $\hat{q} \in \hat{Q}_x$, then

$$\hat{q}(\varphi\{g'(x)(y_1 - y_2) - [g(x + z + y_1) - g(x + z + y_2)]\}) \leqslant \epsilon_0 \hat{q}(y_1 - y_2)$$

where φ is the inverse of the restriction of $g'(x)$ to $\hat{K}_x(g)$;[4]

(b) If h is an element of $K_x(g)$, then $\epsilon > 0$ and $\hat{q} \in \hat{Q}_x$ together imply that there is a $\delta > 0$ such that

$$\hat{q}(\varphi\{g'(x)\, k - [g(x + k + z) - g(x + z)]\}) \leqslant \epsilon \hat{q}(k)$$

if $\hat{q}(k) < \delta$, $z \in U_\epsilon h$, $k \in \hat{K}_x(g)$ and $z + k \in \mathbf{N}$.[5]

The notion of admissibility is essentially an expression of some of the basic requirements of the implicit function theorem.

We suppose that H and E are complete from now on. We are thus ready to prove a theorem which is a form of the Lagrange multiplier rule for minimization problems in locally convex spaces.

Theorem. *Let J be a mapping of H into R and assume that x^* is a local minimum of J on the variety $S(g)$. If $dJ(x^*, \cdot)$ exists and if x^* is admissible for g, then $dJ(x^*, \cdot)$ is an element of the annihilator of $K_{x^*}(g)$* ([5]).

Proof. Let h be an element of $K_{x^*}(g)$ and let $[h]$ denote the subspace of

---

[4] This should be compared with (c8*) of Section 5.

[5] This should be compared with (c8) of Section 5.

H spanned by h. We let $\Sigma = [h] \oplus \hat{K}_{x^*}(g)$ and we consider the mapping Γ of Σ into E defined by

$$\Gamma(ah, k) = g(x^* + ah + k)$$

where $a \in R$ and $k \in \hat{K}_{x^*}(g)$.

Since x^* is admissible for g, it follows from the implicit function theorem that there are, (i) an interval $U_\epsilon \subset R$ with $0 \in \operatorname{int}(U_\epsilon)$, (ii) a neighborhood $\mathbf{N}$ of 0_H, and, (iii) a mapping α of $U_\epsilon h$ into $\hat{K}_{x^*}(g)$ such that $\alpha(0) = 0_H$, $ah + \alpha(a) \in \mathbf{N}$ and $\Gamma(ah, \alpha(a)) = 0_E$ for all a in U_ϵ. Moreover, α is differentiable at 0 and

$$d\alpha(0, a) = -\varphi\{g'(x^*)\, h\}\, a$$

where φ is the inverse of the restriction of $g'(x^*)$ to $\hat{K}_{x^*}(g)$. But h is an element of $K_{x^*}(g)$ so that $g'(x^*)\, h = 0$ and hence, $d\alpha(0, a) = 0_H$. It follows that

$$\alpha(a) = \alpha(0) + d\alpha(0, a) + o(a) = o(a) \tag{*}$$

for all a in U_ϵ.

Now $dJ(x^*, \cdot)$ exists so that

$$J(x^* + ah + \alpha(a)) - J(x^*) = dJ(x^*, ah) + dJ(x^*, \alpha(a)) + o(a)$$

for a in U_ϵ. We claim that $dJ(x^*, \alpha(a))$ is $\mathcal{O}(a)$. Assuming for the moment that this claim is valid, we deduce that

$$J(x^* + ah + \alpha(a)) - J(x^*) = a\, dJ(x^*, h) + o(a)$$

for all a in U_ϵ. Since $x^* + ah + \alpha(a)$ is an element of $S(g)$, since x^* is a local minimum of J on $S(g)$, and since a may be positive or negative as 0 is an interior point of U_ϵ, we conclude that $dJ(x^*, h) = 0$. Thus we need only show that $dJ(x^*, \alpha(a)) = o(a)$. Now $dJ(x^*, \cdot)$ is a continuous linear functional on H and hence, on $\hat{K}_{x^*}(g)$. Thus, there is a $\hat{q}$ in $\hat{Q}_{x^*}$ such that

$$| dJ(x^*, k)| \leqslant \hat{q}(k)$$

for all k in $\hat{K}_{x^*}(g)$. But $\alpha(a) = o(a)$ by (*) so that $\epsilon > 0$ and $\hat{q} \in \hat{Q}_{x^*}$ imply that there is a $\delta > 0$ such that $\hat{q}(\alpha(a)) \leqslant \epsilon \mid a \mid$ if $\mid a \mid < \delta$. It follows that $\epsilon > 0$ implies that there is a $\delta > 0$ such that

$$| dJ(x^*, \alpha(a))| \leqslant \epsilon \mid a \mid$$

if $\mid a \mid < \delta$; i.e., $dJ(x^*, \alpha(a))$ is $o(a)$. The proof of the theorem is now complete.

COROLLARY. *If H is a Hilbert space, then $dJ(x^*, \cdot) = \langle \nabla J(x^*), \cdot \rangle$ with $\nabla J(x^*)$ orthogonal to $K_{x^*}(g)$ and hence, there are real numbers λ_β such that $\nabla J(x^*) = \Sigma \lambda_\beta e_\beta$ where $\{e_\beta\}$ is an orthonormal basis of $\hat{K}_{x^*}(g)$.*

REFERENCES

1. FALB, P., *Direct Methods in Optimal Control,* McGraw-Hill, New York (to appear in 1968).
2. LIUSTERNIK, L. AND SOBOLEV, V., "Elements of Functional Analysis." Ungar, New York, 1961.
3. MICHAL, A., Differential calculus in linear topological spaces. *Proc. Natl. Acad. Sci. U. S.* **24** (1938), 340–342.
4. HYERS, D., A generalization of Frechet's differential. *Rev. Cienc.*, Lima (Peru), **47** (1945), 645–663.
5. HORVATH, J., "Topological Vector Spaces and Distributions." Addison-Wesley, Reading, Massachusetts, 1966.
6. DIEUDONNÉ, J., "Foundations of Modern Analysis." Academic Press, New York, 1960.
7. AVERBUCK, V. I., AND SMOLYAKOV, O. G., Differentiability in linear topological spaces. *Dokl. Akad. Nauk SSSR*, T. 173, 1967.
8. BASTIANI, A., Applications différentiables et variétés différentiables de dimension infinie. *J. Analyse Math.* **13** (1964), 1–114.
9. FRÖLICHER, A., AND BUCHER, W., Calculus in vector spaces without norm. *In* "Lecture Notes in Mathematics," No. 30, Springer, Berlin, 1966.